"Assimilation of TOPEX/POSEIDON Data into an Ocean General Circulation Model"

Ichiro Fukumori

Jet Propulsion Laboratory
California Institute of Technology

Abstract

The nature of global, large-scale sea level variability is studied using a numerical ocean circulation model. Salient physics underlying sea level change is identified and used to analyze TOPEX/POSEIDON measurements by data assimilation. The study demonstrates how and what can be resolved and improved by assimilation. Proper evaluation of errors, particularly model representation error is critical in reconciling observations and model estimate.

1. Introduction

Sea level fluctuates in time reflecting not only processes of the sea surface but also the state of the ocean at depths, making satellite altimetry a powerful tool for studying global ocean dynamics and thermodynamics. Yet, direct inferences of subsurface flow and density from sea level is not straightforward. This study examines the physical nature of large-scale sea level variabilities and the underlying circulation by a combined analysis of sea level data from TOPEX/POSEIDON and a numerical ocean circulation model.

Estimating ocean circulation from sea level data is in essence an inverse problem, whose forward solution is provided by numerical models of the ocean. Realistic numerical simulations of the ocean are now commonplace, and are useful in dynamically interpreting observations, and in making inferences of the underlying physical processes. More recently, such models are employed in synthesizing data, i.e., data assimilation, which allows direct estimation of the entire ocean circulation from partial measurements of the oceanic state.

The problem of what changes in large-scale circulation are reflected in sea level is first examined in Section 2, by analyzing results of an ocean general circulation model. Sea level variability from a wind- and thermally-driven numerical simulation of the global ocean is decomposed into its dependence on forcing and the nature of the ocean's response. These results are then employed in Section 3 to assimilate TOPEX/POSEIDON data into the model, so as to identify and extract changes in large-scale ocean circulation measured by the altimetric satellite. One of the foci of this study is to explore and demonstrate how and what can be inferred of ocean circulation from assimilating actual satellite altimetric measurements.

2. Nature of Sea Level in Relation to Atmospheric Forcing

The correspondence between sea level changes and ocean circulation is examined by analyzing a numerical simulation of the global ocean. Here, the sea level signal of interest is those apart from tides and effects of atmospheric pressure loading, and also exclude signals associated with mesoscale variabilities, which, in fact, are the three largest signals measured by an altimeter. Instead, the focus of the study is on the remaining large-scale changes and their corresponding circulation.

The model employed is based on the Modular Ocean Model of NOAA's GFDL. The model domain is global, from 80°S to 80°N, with a coarse 2° by 1° horizontal grid with 12 vertical layers ranging in thickness from 50 m near the surface to 1000 m at depth. The model is run from January 1992 till December 1993 forced by daily winds from the NCEP atmospheric analyses and COADS climatological monthly heat flux estimates, following an eight-year spin-up using stationary forcings.

Plate 1 summarizes the nature of the model's sea level changes as a function of space and its physics. The overall root-mean-square (rms) amplitude (Plate 1a) varies spatially with some areas exceeding 8 cm. On average, the total sea level variability is on the order of 5 cm rms. Effects of heating and cooling can be isolated (Plate 1b) by running the model with winds held stationary as during the spin-up phase but using monthly heat flux forcing as in the control run (Plate 1a). Comparisons between Plates 1a and 1b, demonstrate seasonal heating and cooling effects dominating sea level variability in mid-latitudes. A closer examination shows such thermal forcing only locally affecting temperature of the model's upper most levels. Effects of heating and cooling is small in the tropics because of little seasonal change in heat flux. Thermal effects at higher latitudes are negligible due to small thermal expansion coefficients reflecting lower sea surface temperatures.

Differences between Plates 1a and 1b are due to wind forcing, which can be decomposed into contributions from baroclinic (Plate 1c) and barotropic (Plate 1d) changes in circulation. Effects of winds are isolated by running the model with daily wind forcing but with stationary thermal forcing. The decomposition is achieved by computing the two parts separately [e.g, *Pinardi et al.*, 1995]. Comparisons among different panels in Plate 1 demonstrate that such decompositions are nearly linear, in spite of model nonlinearities.

Baroclinic changes dominate low latitudes, whereas barotropic changes dominate variabilities at high latitudes. The baroclinic response is largest in the tropics owing to spectral characteristics of baroclinic planetary waves. Namely, the ocean's response to winds depends in part on properties of oceanic waves that carry energy away from forcing regions. The maximum resolvable frequency is largest at the equator, and as a result has the largest spectral window (and wind energy) available for baroclinic response. Although wind changes are larger at higher latitudes, they occur at frequencies outside the domain of baroclinic waves. The baroclinic response can further be decomposed into separate dynamic modes, where the first mode contribution is found to be largest in most locations.

Barotropic waves, on the other hand, have significantly higher frequencies than baroclinic waves, covering most of the frequency range of the winds. As a result, the barotropic response does not depend as much on the winds' frequency dependence as do baroclinic waves, and is largest at high latitudes where wind variability is largest.

The dominance of barotropic response explains the high frequency nature of sea level observed at high latitudes (e.g., Figure 1a), whereas sea level records in the tropics are characterized by low frequencies (periods longer than 100-days). In fact, many high latitude regions of the model have half their intra-seasonal sea level variabilities at periods shorter than twenty days. On the other hand, the largest barotropic response, found in the Bellingshausen abyssal plain, has a dominating period of about 30-days (Figure 1b). The large barotropic response in the Bellingshausen basin apparently reflects presence of a barotropic resonance locally forced by winds, due to closed potential vorticity contours surrounding the area [e.g., *Willebrand et al.*, 1980]. Figure 1b demonstrates that a similar oscillation is observed in TOPEX/POSEIDON data, albeit with a smaller amplitude.

3. Assimilation of TOPEX/POSEIDON Data

Section 2 analyzed the forward relationship between ocean circulation and sea level. This section examines the inverse; Namely, ocean circulation underlying measured sea level changes of TOPEX/POSEIDON is estimated by assimilating the altimetric data into the same model used in Section 2.

Three years of TOPEX/POSEIDON data from January 1993 until December 1995 was assimilated into the model to extract and extrapolate large-scale sea level signals in the measurements and to explore how the model circulation can be improved. All standard environmental corrections are applied to the data. Time-continuous, along-track sea level anomalies are directly used in the analysis as opposed to space-time maps of the data, so as to avoid possible aliasing of high frequency barotropic signals at high latitudes (cf. Section 2).

The assimilation is performed using an approximate Kalman filter and smoother, which involves approximating the state's error covariance matrix by a reduction of its effective dimension and using its time-asymptotic limit [*Fukumori et al.*, 1993, 1995].

For instance, the analysis in the previous section showed that there are three dominant modes underlying large-scale sea level changes. They are wind-driven barotropic and first baroclinic modes, and seasonal heating and cooling effects of the upper most layers. However, since heating and cooling effects are local and have relatively small dynamic effects on the present time-scales of interest, the physics of large-scale sea level changes may effectively be approximated in terms of just barotropic and first baroclinic modes as opposed to the model's 12 vertical degrees of freedom. The horizontal reduction is achieved by mapping the two dynamic mode amplitudes onto a coarser grid than that of the forward model, thereby isolating and defining the large-scale. The coarse grid is defined on a 10° by 5° zonal and meridional grid, respectively, and the transformation onto the model grid is achieved by objective interpolation [e.g., *Bretherton et al.*, 1976].

3.1 Error Identification and Calibration

Data assimilation can be defined mathematically as an inverse problem. A least-squares estimate ($\mathbf{x}$) is sought that minimizes the weighted sum of differences (J) with data ($\mathbf{y}$) and prior model estimate ($\mathbf{x}_0$).

$$J = \Sigma_t(\mathbf{y} - \mathbf{y}(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{y}(\mathbf{x})) + \Sigma_t(\mathbf{x} - \mathbf{x}_0)^T\mathbf{P}^{-1}(\mathbf{x} - \mathbf{x}_0) \tag{1}$$

The summation is taken over all available data and time, and weights $\mathbf{R}$ and $\mathbf{P}$ are error covariance matrices of data and model, respectively [e.g., *Wunsch*, 1997]. While the solution of Eq (1) is mathematically straightforward (e.g., Kalman filter), specification of $\mathbf{R}$ and $\mathbf{P}$ requires care; A mis-specification of these weights amounts to solving a different problem.

Cohn [1997] has a clearer than usual description of the exact nature of these errors. In particular, data error not only corresponds to errors of the measuring instrument, but also what is called model represetation error. Namely, observations can be written as a function ($\mathbf{E}$) of the true oceanic state, $\mathbf{w}$, plus instrument error, ϵ;

$$\mathbf{y} = \mathbf{E}(\mathbf{w}) + \epsilon \tag{2}$$

Instrument errors represent quantities unrelated to either model or ocean. For satellite altimetry, ϵ includes, for example, errors in the satellite's orbit and ionospheric corrections. Function $\mathbf{E}$ represents the measurements' sampling operation.

In terms of quantities in model space, Eq (2) can be rewritten as;

$$\mathbf{y} = \mathbf{H}(\hat{\mathbf{x}}) + \{\mathbf{E}(\mathbf{w}) - \mathbf{H}(\Pi\mathbf{w})\} + \epsilon \tag{3}$$

where $\mathbf{H}$ is the model equivalent of $\mathbf{E}$, $\hat{\mathbf{x}}$ is the true model state, and operator Π projects the infinite dimensional oceanic state to the model space (i.e., $\Pi\mathbf{w} \equiv \hat{\mathbf{x}}$). Assimilation is the inversion of (3) that relates model state to observations rather than a solution of (2). The second term in (3) describes differences between the ocean and the finite dimension of the model, and, as far as the model is concerned, is indistinguishable from errors of the measuring instrument. The second term in (3) is the representation error, and corresponds to physics missing from the model but represented in the observation. Representation errors are inconsistent with model physics and are typically quantities whose scales are smaller than model grid spacing. In effect, representation errors downweight the data constraint (first term in Eq 1) and prevent models from being forced too close to observations that it cannot represent, thus guarding against model 'indigestion'.

Fu et al., [1993] introduced an objective means of estimating data and model errors based on a comparison between observations and model simulation. Plate 2 shows diagonal elements of $\mathbf{R}$ and $\mathbf{P}$ (projected to sea level) based on such estimate. Data error is larger than model error estimate, and in fact much larger than the measurement accuracy of TOPEX/POSEIDON ($\approx$ 2 cm rms). This is due to representation error dominating data error, in particular, sea level associated with meso-scale eddies.

Namely, the largest variability in the corrected altimeter data is due to meso-scale eddies, which are spatially smaller than what the present model can resolve. In fact, the spatial dependence of data error variance closely resembles distribution of meso-scale variability.

Model process noise (incremental error of the model) was modeled in the form of wind error, and calibrated such that the resulting simulation error estimate based on the Kalman filter algorithm (Plate 2c) is comparable with estimates obtained from the data–simulation comparison (Plate 2b). Such objective means of identifying and calibrating data and model errors assures consistency and accuracy of the assimilation as described below.

3.2 Validation of Estimate

As in most inverse problems, altimetric assimilation is rank deficient and has an infinite number of solutions that can reduce model–data misfits. That particular results are reasonable and accurate can be assessed by examining self-consistencies of the calculation and by comparisons with independent observations.

Consistency of the calculation is evaluated by comparing assimilated results against error estimates that were assumed. For instance, Plate 3a shows improvements in the filtered estimate's sea level from that of the simulation in terms of reduction of model–data misfit, $\mathbf{y} - \mathbf{y}(\mathbf{x})$. Positive values indicate assimilation's improvement, which is the case almost everywhere. Plate 3b is an independent estimate of the same quantity based on formal error estimates used in the assimilation. The qualitative and quantitative similarities between Plates 3a and 3b demonstrate the statistical consistency of the present assimilation.

Theoretically, model errors are nonincreasing functions of the amount of data that is assimilated. Therefore, if assimilation is done correctly, the assimilated estimates should be more accurate than that of the simulation, regardless of property. Figure 2 demonstrates examples of comparing model estimates (assimilation and simulation) with independent in situ measurements; The examples are subsurface temperature, current, and bottom pressure measurements. In each case, the altimetric assimilation is in better agreement with in situ measurements, consistent with formal errors that are also shown in the figures. Remaining discrepancies are partly due to limitations in the information content of the data. For instance, although resolving lower frequencies, the altimetric assimilation was not able to correct high frequency errors of bottom pressure (periods shorter than 10-days in Fig 2c) in spite of the model's coherence with the data. This is likely in part due to TOPEX/POSEIDON's repeat cycle being 10-days, which is not sufficient to correct such high frequency variations.

Most comparisons demonstrate similar improvements where altimetric assimilation is in better agreement with independent in situ measurements than model simulation is. On the other hand, a few comparisons result in larger discrepancies after the assimilation. However, such situations are found to be due to representation error dominating the measurements rather than a failure of the assimilation per se. Figures 3a and 3b compare model estimates with subsurface temperature at the same location but at different depths, and is an example having larger model–data differences after assimilation. In this example, the dominant variability in the data is incoherent with depth and is something the model is qualitatively lacking, indicative of model representation error.

3.3 Circulation Estimates

Model corrections by the assimilation are mostly baroclinic in nature in the tropics but barotropic at higher latitudes, similar to properties of model sea level itself (Section 2). Plate 4 shows an example of how model circulation is modified by the assimilation at 1200m depth for a particular day in October 1993. A substantial velocity correction is found in the Bellingshausen abyssal plain without associated temperature corrections, indicative of barotropic change. Such modification reflects the assimilation successfully reducing the excessive amplitude of the model's barotropic resonance as was evident in Fig 1. The excess resonance is apparently due to the model's coarse vertical resolution that results in too large a horizontal gradient of potential vorticity and thereby the magnitude of the resonance.

Substantial temperature and circulation changes are also found in other regions of Plate 4, especially the Indian Ocean. A property of particular interest is the meridional heat flux. Figure 4 shows zonally and vertically integrated time-averaged meridional heat flux for the Indian, Pacific, and Atlantic Oceans as a function of latitude. Because of the filter's linear nature and that the assimilated data are anomalies with respect to the three-year mean, the model's time-mean circulation is not directly constrained and the mean of the assimilated estimate is hardly different from that of the simulation. The discontinuity in the Pacific about 2°S (and in the Indian Ocean between 6°S and 10°S) is due to the Indonesian Throughflow which carries approximately 0.6 PW of heat from the Pacific into the Indian Ocean. The mean flux estimates are somewhat smaller than other estimates because of the coarse nature of the model, which necessarily smooths over and therefore misses the narrow boundary currents where the temperature and velocity correlations are high.

Temporal variability of meridional heat flux is largest in the tropics, where seasonal changes are larger than even the mean (Fig 4). Such variations are largely due to fluctuations in surface layers where mean temperature and velocity fluctuations are largest. Northward meridional heat flux is maximum during boreal winter months and minimal in summer in all three basins. The largest seasonal cycle occurs in the Indian Ocean associated with the monsoon, where the North Equatorial Current ($\sim$ 5°N) disappears during the summer replaced with southward flow throughout the interior of the tropical Indian Ocean feeding the South Equatorial Current (10 $\sim$ 20°S). The Indonesian Throughflow is also maximum during boreal summer further enhancing the southwestward transport of the South Equatorial Current.

Modifications in heat flux estimates due to assimilation is much smaller than the model's temporal variability, but is largest in the tropics as in the model fluxes themselves. Much of the modifications occur at high frequencies, but nontrivial changes are also found at lower frequencies especially the annual and semi-annual periods. Examples of such differences are shown in Figure 5 where the changes are largest. For instance, the simulation's annual component is attenuated at 6°S in the Indian ocean. In comparison, the assimilation has greatly strengthened the variability along 4°S in the Atlantic Ocean, where the model simulation significantly underestimated the temporal change. The example along 2°N in the Pacific, on the other hand, shows the phase of the annual cycle being altered (advanced) by the altimetric assimilation while the amplitude is hardly different. The changes by assimilation are almost 180° out of phase between Figs 5a and 5b, in part, due to modifications of the Indonesian Throughflow.

4. Conclusion

A latitudinal delineation exists in the nature of large-scale global sea level variability due to inhomogeneities in ocean physics; Wind-driven baroclinic and barotropic changes dominate the tropics and high latitudes, respectively, while seasonal heating and cooling effects control sea level at mid-latitudes. Identification of such characteristics facilitate analyses of altimetric data and were taken advantage of in assimilating TOPEX/POSEIDON data into an ocean circulation model.

Creating a comprehensive description of the ocean is becoming increasingly recognized as an important and urgent task in understanding the turbulent nature of the ocean and its role in climate fluctuations. In particular, ocean data assimilation has received much attention in recent years as a means of synthesizing diverse observations into a physically consistent estimate of the oceanic state. While much of the focus in assimilation has concerned methodologies, a more fundamental issue is the identification of what exactly is being solved.

The present study illustrates how the Kalman filter approach, together with an analysis of the forward model, provides a quantitative framework to properly set up the assimilation problem, and to solve it consistently. In particular, evaluation of model representation error is found to be critical, as it identifies the model range space and because it defines the assimilation problem itself. The analysis demonstrates consistent and quantitative improvements of the model. Such results assures the accuracy of the estimate, and therefore provides a basis to further understanding of processes controlling changes in ocean circulation.

Acknowledgment

This research was carried out in part by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

Bretherton, F. P., R. E. Davis, and C. B. Fandry, 1976. A technique for objective analysis and design of oceanographic experiments applied to MODE–73, *Deep-Sea Research*, **23**, 559–582.

Cohn, S. E., 1997. An introduction to estimation theory, *Journal of the Meteorological Society of Japan*, **75**, 257–288.

Fu, L.-L., I. Fukumori and R. N. Miller, 1993. Fitting dynamic models to the Geosat sea level observations in the Tropical Pacific Ocean. Part II: A linear, wind-driven model, *Journal of Physical Oceanography*, **23**, 2162–2181.

Fukumori, I., J. Benveniste, C. Wunsch and D. B. Haidvogel, 1993. Assimilation of sea surface topography into an ocean circulation model using a steady-state smoother, *Journal of Physical Oceanography*, **23**, 1831–1855.

Fukumori, I., and P. Malanotte-Rizzoli, 1995. An approximate Kalman filter for ocean data assimilation; An example with an idealized Gulf Stream model, *Journal of Geophysical Research*, **100**, 6777–6793.

Pinardi, N., A. Rosati, and R. C. Pacanowski, 1995. The sea surface pressure formulation of rigid lid models. Implications for altimetric data assimilation studies, *Journal of Marine Systems*, **6**, 109–119.

Willebrand, J., S. G. H. Philander, and R. C. Pacanowski, 1980. The oceanic response to large-scale atmospheric disturbances, *Journal of Physical Oceanography*, **10**, 411–429.

Wunsch, C., 1996. "The Ocean Circulation Inverse Problem", Cambridge Univ. Press, New York, NY, 442pp.

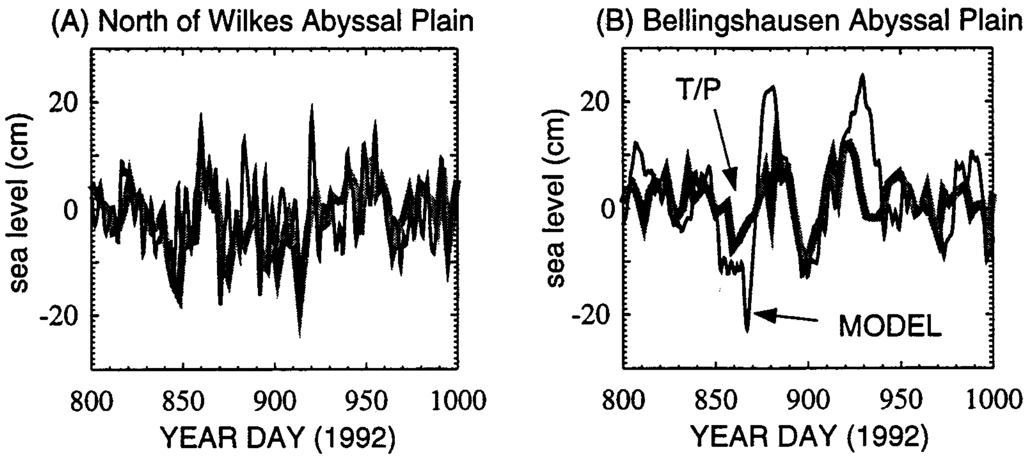


Figure 1: Model (black) and T/P (gray) sea level comparison; (A) 100°E, 55°S (north of Wilkes Abyssal Plain), (B) 95°W, 60°S (Bellingshausen Abyssal Plain). The T/P data denote 3-day averages within the vicinity of the respective points. The model is sampled at 1-day intervals.

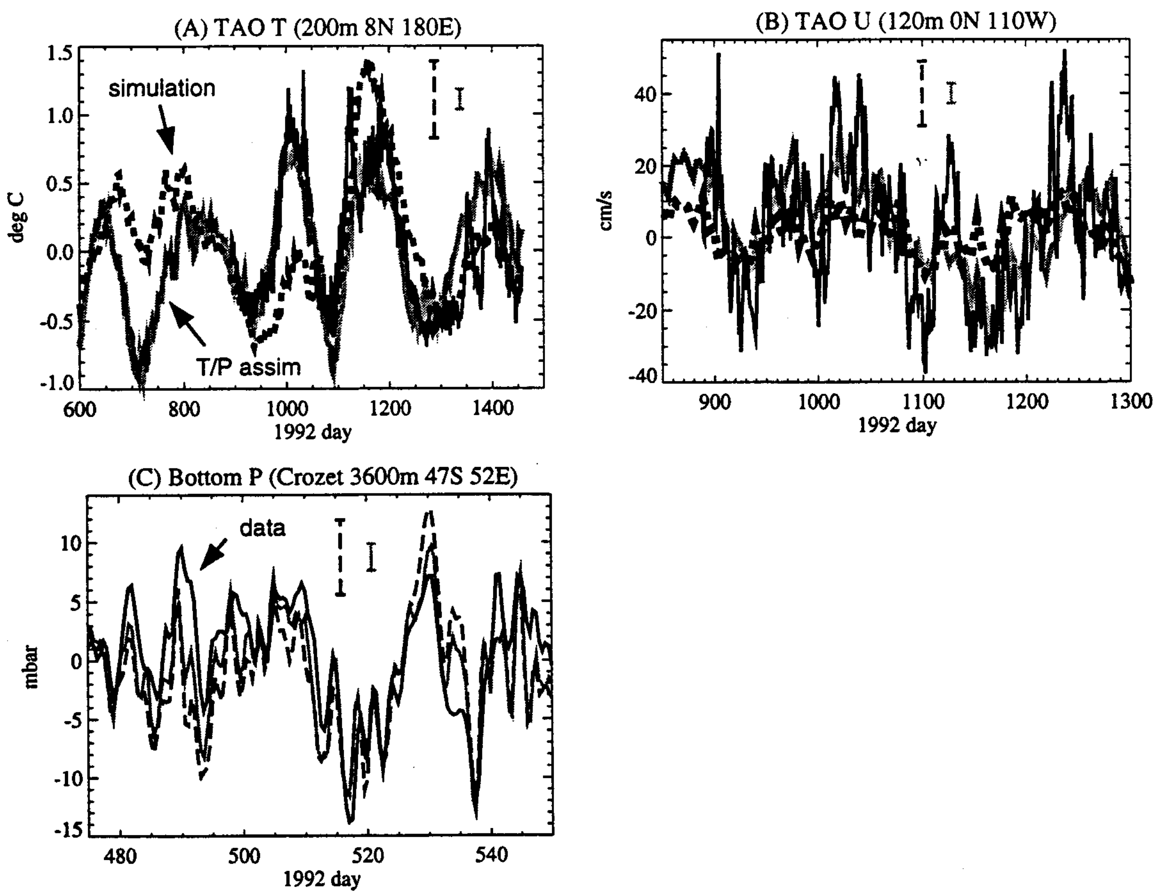


Figure 2: Comparisons of T/P assimilation (gray) and model simulation (dashed black) with in situ measurements (solid black); (A) temperature (200 m, 8°N, 180°E; TAO), (B) zonal velocity (120m, 0°N, 110°W), (C) bottom pressure (3600m, 47°S, 52°E). The bars denote standard error estimates.

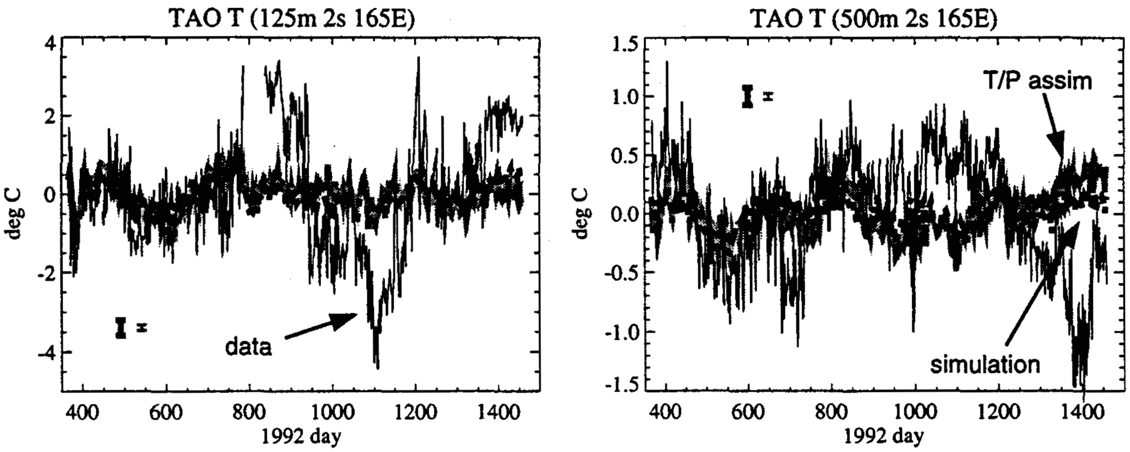


Figure 3: Comparisons of T/P assimilation (gray) and model simulation (thick dashed black) with temperature measurements (thin solid black) at 2°S 165°E; (A) 125 m, (B) 500 m. The bars denote standard error estimates.

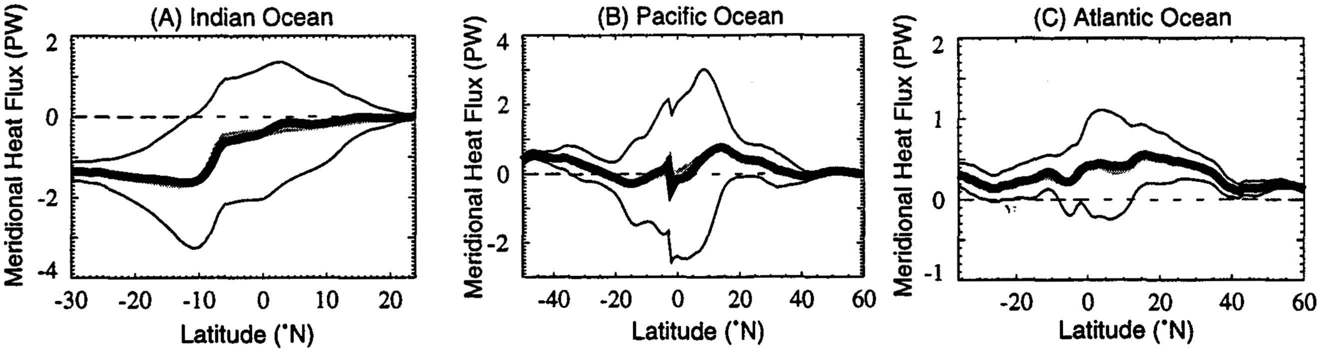


Figure 4: Zonally and vertically integrated meridional heat flux (PW); (A) Indian Ocean, (B) Pacific Ocean, (C) Atlantic Ocean. The time-mean of T/P assimilation (thick black) and model simulation (thick gray) are practically identical. The thin solid curves are mean assimilation flux plus minus root-mean-square of 30-day averaged fluxes.

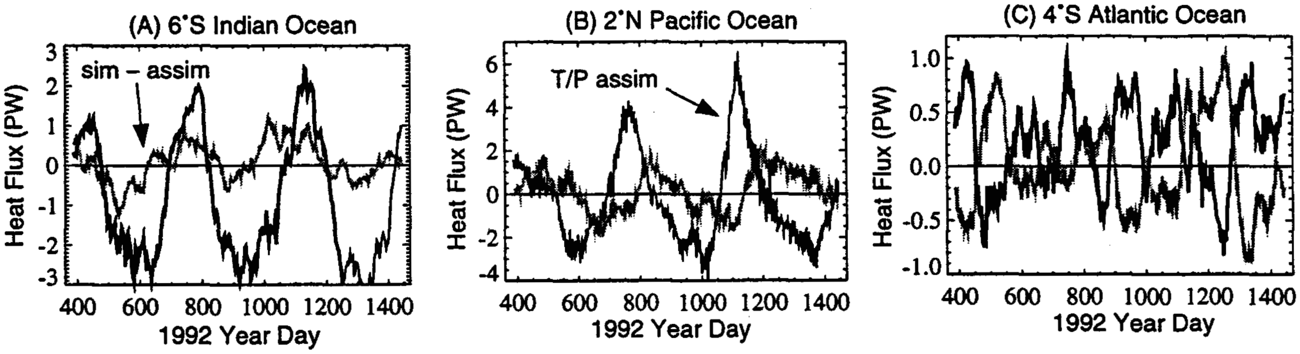


Figure 5: Examples of zonally and vertically integrated heat flux (PW) as a function of time; (A) 6°S in the Indian Ocean, (B) 2°N in the Pacific, and (C) 4°S in the Atlantic. The figures show the T/P assimilated estimate (black) and the difference of simulation minus assimilation (gray).

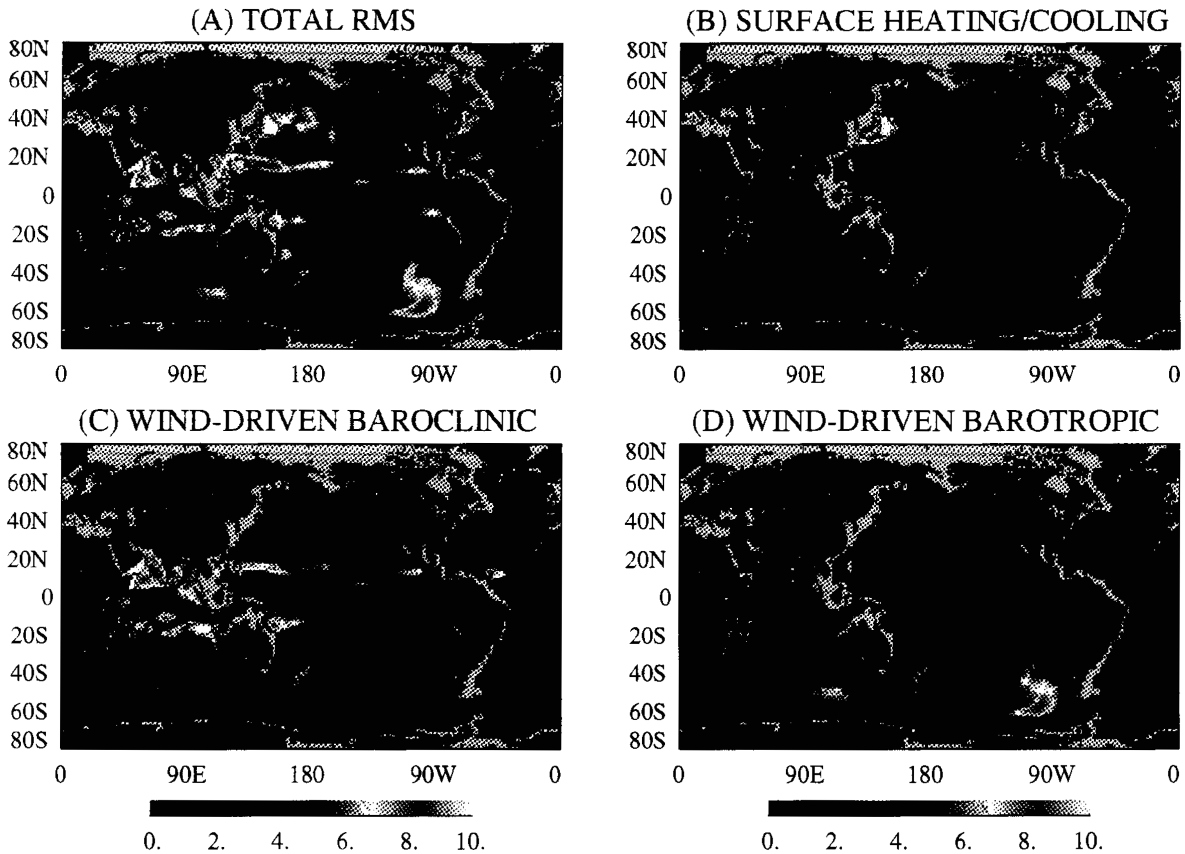
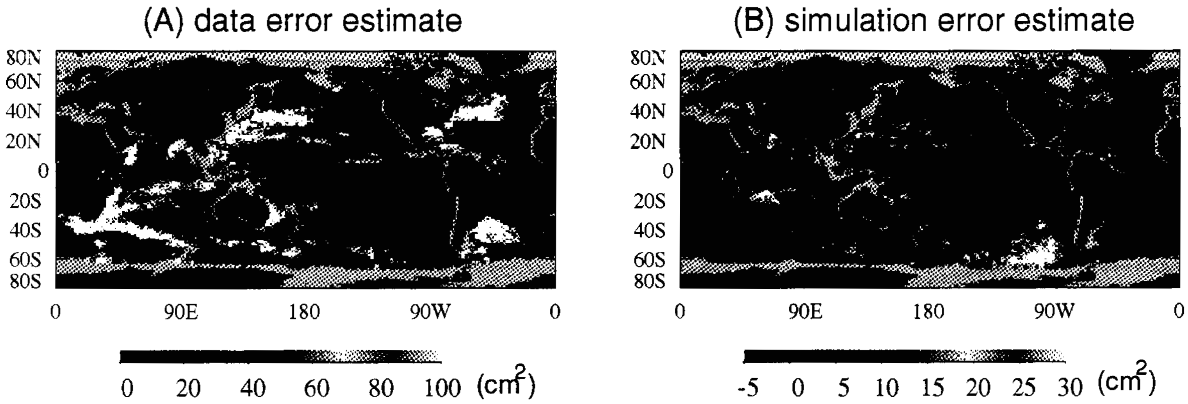


Plate 1: Nature of sea level variability simulated by an OGCM. Sea level variability (root-mean-square; units cm) due to wind and thermal forcing shown in (A) is separated into parts associated with (B) surface heating/cooling, (C) wind-driven baroclinic, and (D) wind-driven barotropic motion.



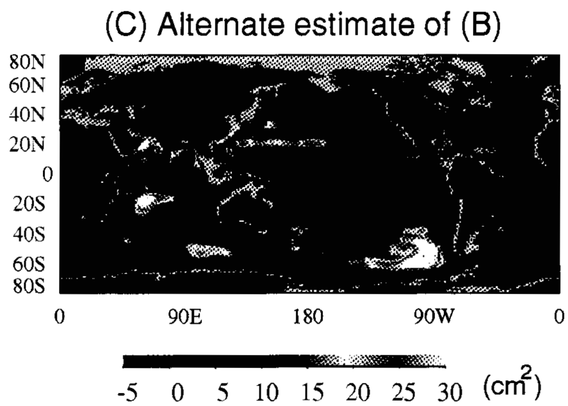


Plate 2: Error Calibration. (A) and (B) are data and model simulation error estimates, respectively, in terms of sea level variance, based on a data/simulation comparison. (C) is same as (B), but from the Kalman filter algorithm using scaled wind-stress as model process noise.

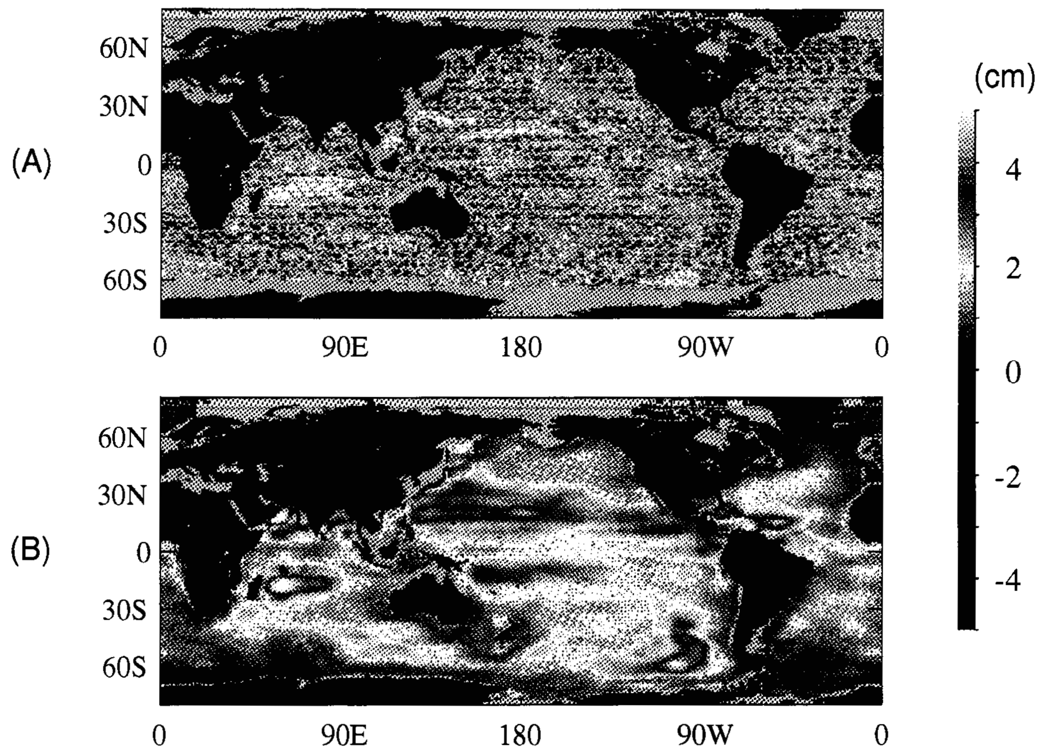


Plate 3: Model improvement of simulating observed T/P sea level variability by assimilation. Values are reduction of root-mean-square sa level residual; (A) Skill of filtered forecast relative to simulation, (B) expected value of (A). Positive (negative) numbers indicate improvements over (worsening from) simulation.

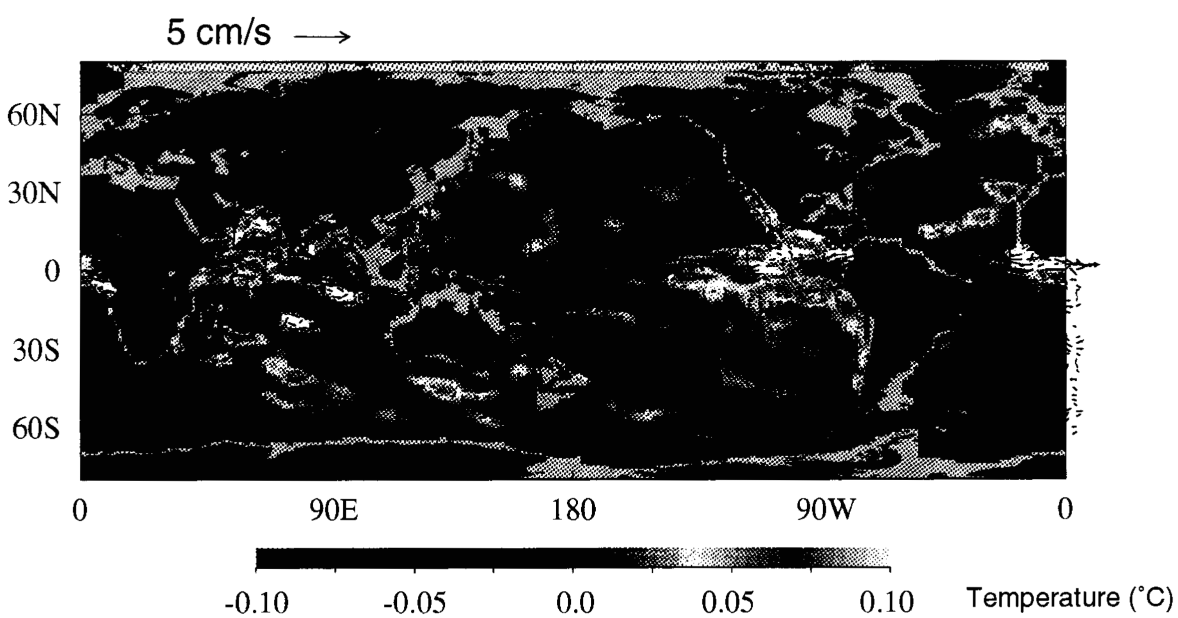


Plate 4: Differences between T/P assimilation and model simulation at 1200 m for a particular instant in October 1993. Colors and arrows denote temperature and velocity differences, respectively.